

1 **USER INFORMATION COORDINATION ACROSS MULTIPLE DOMAINS**

2 **FIELD OF THE INVENTION**

3 This invention is directed to the field of computer
4 networks. It is more particularly directed to the
5 Internet, trackers and servers that use cookies.

6 **BACKGROUND OF THE INVENTION**

7 The Internet Protocol (usually referred to as IP)
8 provides network connectivity to users across the
9 world. The most common application in networks running
10 this protocol is the HTTP protocol, which allows a
11 web-browser to access a web-server over the Internet.
12 HTTP is a request-response protocol, and is designed to
13 be stateless. A stateless protocol is one that does not
14 require either the client or server to remember any
15 information from prior interactions.

16 For many types of web-based exchanges over the
17 Internet, it is desirable to maintain some state across
18 the different requests of the HTTP protocol. We refer
19 to a scheme that can identify an user across multiple
20 HTTP sessions as an user tracking mechanism. The most
21 common user tracking mechanisms is for the web-server
22 to store a cookie at the web-browser. A cookie is data
23 that is placed within the web-browser by a client. This
24 data is sent to the server by the browser whenever it

1 makes a new request to the browser. Typically cookies
2 are used to store the identity of an user so that
3 multiple visits can be correlated. They can also store
4 the profile or preferences of an user, or security
5 credentials which allow an user to access specific
6 content at a web-server.

7
8 When a server places a cookie on the browser, it can
9 specify that the cookie be sent to servers other than
10 itself. Adding other sites to the site to which the
11 cookies can be sent allows cookie information to be
12 shared with other servers. Restricting the sites that a
13 cookie gets delivered helps in maintaining the security
14 and privacy of data placed in the cookies. However, the
15 current implementation of cookies in web browsers
16 restricts the set of servers that can be specified to
17 receive the cookie set in this manner. If a server sets
18 a cookie, it can also request that the cookie be sent
19 to other servers which share a domain name suffix with
20 it. Thus, a server with domain name,

21 www.watson.ibm.com

22 can set a cookie to be set in the browser so that the
23 cookie is sent only to

24 www.watson.ibm.com,

25 or to any machine with the name ending in

26 watson.ibm.com,

27 or to any machine with the name ending in

28 ibm.com,

29 or to any machine with the name ending in '.com'. The
30 last choice in the list will send the cookie to all the
31 machines in the '.com' domain. If a cookie contains
32 information that is sensitive, e.g. the security

1 credentials of the users, it is highly undesirable that
2 the information be sent to many machines.

3 In many situations, it is desirable that the cookie
4 information be shared with members of another domain .
5 As an example, a server

6 www.watson.ibm.com
7 may want to share its cookie information with the
8 server,

9 www.berkeley.edu.

10 However, with the current way cookies are supported
11 does not make it possible to set a cookie which will
12 only be shared between these two servers. The only
13 option would be to have a cookie that is sent to all
14 the servers within the Internet, which is highly
15 undesirable.

16 The same problem is experienced by other user tracking
17 mechanisms. As an example, one common user tracking
18 mechanism uses URL rewriting. In this mechanism, the
19 content presented to an user is rewritten so that an
20 unique tag is present in all links that the user may
21 access. As the user clicks on the appropriate link, the
22 tag is carried on to the site, and identifies the user
23 across the sessions. When two sites use independent
24 tags to track users, they are unable to correlate the
25 user at one site with the user on the other site.

26 For purposes of this invention, we use the term user
27 tracking mechanisms to refer to cookies; URL rewriting
28 or other techniques that are used to identify users
29 accessing a web-site; a domain to refer to a set of

1 servers with whom the normal operation of the
2 user-tracking mechanism can be used to share
3 operations; and an user tracker as a server which
4 employs an user tracking mechanism. It would be
5 advantageous to be able to use the same user tracking
6 mechanism across more than one domain, in which
7 heretofore normal operation of the user tracking
8 mechanism can not be used.

9 **SUMMARY OF THE INVENTION**

10 It is therefore an aspect of the present invention to
11 provide a method by which two web servers and/or user
12 trackers operating in two different domains can
13 correlate user tracking information.

14 It is a further aspect of the invention to provide an
15 apparatus by which two servers and/or user trackers
16 operating in two different domains can correlate user
17 tracking information.

18 It is a further aspect of the invention to enable a
19 same user tracking mechanism to be used across more
20 than one domain, where normal operation of the user
21 tracking mechanism can not be used.

22 It is a further aspect of the present invention to
23 provide a method and apparatus by which two web-servers
24 and/or user trackers operating in two different domains
25 can correlate cookies placed into a browser
26 independently by them.

1 It is a further aspect of the present invention to
2 provide a method and apparatus by which two web-servers
3 and/or user trackers can correlate user tracking
4 information created as a result of URL rewriting
5 mechanisms.

6 In an example embodiment of the invention, a web server
7 and/or user trackers in one DNS domain establishes a
8 cookie containing an identity field at a client's
9 browser, redirects the client to a second web-browser
10 with an uRL containing the identity field created in
11 the cookie. The second web-browser creates a cookie
12 with a second identity field, and stores the first
13 identity field and the second identity fields in a
14 global database. The database information is retrieved
15 by the two web-servers to correlate the cookie
16 information.

17 In an alternative embodiment, a global database need
18 not be maintained, but rather each web-server maintains
19 its own local database containing the identity of the
20 different users. Each of the servers creates an unique
21 identity for the client browser, and redirects the
22 client to access an uRL at the other server which is
23 used to create a local database correlating the two
24 identities. Links from one server's pages to another
25 are rewritten to carry the unique identities in the two
26 sites. Applications of this invention include, but
27 are not limited to: systems that correlate user
28 identities across multiple domains, systems that
29 provide single sign on support across multiple domains,

1 systems that store user preferences based on client
2 identity, etc.

3 **BRIEF DESCRIPTION OF THE DRAWINGS**

4 These and other aspects, features, and advantages of
5 the present invention will become apparent upon further
6 consideration of the following detailed description of
7 the invention when read in conjunction with the drawing
8 figures, in which:

9 Fig. 1 shows an example of an environment having
10 multiple Internet domains and the problems associated
11 with using cookies established in one domain with those
12 of other domains;

13 Fig. 2 shows an example of a system that would allow a
14 sharing of user information across two or more DNS
15 domains by a web-server;

16 Fig. 3 shows a flowchart that illustrates an example of
17 a method used for sharing user information across two
18 domains by one of the web-servers among a pair of
19 web-servers that wishes to share its user information;

20 Fig. 4 shows a flowchart that illustrates an example of
21 a method used for sharing user information across two
22 domains by the second web-server among the pair that
23 wishes to share their user information; and

1 Fig. 5 shows an example of an apparatus that can be
2 used for sharing user information across web servers
3 that are located in two different domains.

4 Other objectives and a better understanding of the
5 invention may be realized by referring to the detailed
6 description.

7 DESCRIPTION OF THE INVENTION

8 The present invention provides methods and apparatus
9 for sharing cookies and/or cookie-like objects within
10 the Internet, trackers and/or servers. A typical
11 environment in which user information is tracked within
12 an IP network is shown in Fig. 1. It shows a browser
13 101 and three servers 103 105 107. The browsers and the
14 servers are connected over an IP network 113. An
15 example of the IP network 109 would be the public
16 Internet. The IP network consists of several domains,
17 two of which are shown in the figure. The domain 109
18 consists of all servers with the name,

19 domain1.com

20 and it contains two of the servers shown, namely the
21 server,

22 server1.domain1.com

23 103 and,

24 server2.domain1.com

25 105. The domain 111 consists of all servers with the
26 name,

27 domain2.com

28 and it contains the server,

1 server3.domain2.com
2 107. The server and domain names used in the figure are
3 for illustrative purposes only.

4 Within the environment shown in Figure 1, the servers
5 may use a cookie mechanism to track user information.
6 When,

7 server1.domain1.com
8 103 places a cookie on the browser 101, it can instruct
9 that the cookie be shared with the other servers in the
10 domain,

11 domain1.com

12 109. Thus, the two servers 103 and 105 can access the
13 cookies placed into the browser by each other and can
14 track user information by using a shared format for
15 cookie data. However,

16 server1.domain1.com

17 103 can not request that the browser send the same
18 cookie to a server in the other domain

19 domain2.com

20 111. Thus, the cookie information placed on the browser
21 by,

22 server1.domain1.com

23 103 can not be shared by ,

24 server3.domain2.com

25 107 since it is in another domain 111. Under the
26 well-known rules of cookie sharing, the only way such
27 sharing can be obtained is by defining a cookie to go
28 to all machines with a name suffix of '.com'. Clearly,
29 this would be highly undesirable.

1 Instead of cookies, an alternative way to share user
2 identity is to use the technique of URL rewriting in
3 accordance with the present invention. In the context
4 of URL rewriting, an unique identity is assigned to an
5 user when the user first contacts a server. This
6 identity is embedded in the URL which is passed to the
7 user, and all links provided to the user are included
8 in a similar fashion. The identity being used for an
9 user is local to a server. In general, two servers can
10 not share the information about a rewritten URL without
11 explicit prior agreement. As opposed to cookies, the
12 identity association of the user is not stored by the
13 browser, and each identity association is specific to a
14 particular session.

15 As an example of URL rewriting, consider a company
16 which is accessed through its portal

17 `http://www.company.com`

18 The technique of user tracking using URL rewriting
19 would have the web-server for the site redirecting
20 users accessing the site

21 `http://www.company.com`

22 to another URL

23 `http://www.company.com/<identity>/index.html.`

24 The <identity> field is generated as an unique
25 identifier for the specific session. If the links
26 embedded in the page index.html (and other pages) are
27 all relative, or if the server modifies the contents of
28 a page to include the <identity> tag in all referenced
29 links; the <identity> field would be part of the URL
30 whenever the user clicks on any embedded links within
31 the page under the normal conventions of HTTP protocol

1 . By looking at the <identity> field, the web site can
2 determine who the user accessing a page is. However, if
3 the user accesses the page,
4 http://www.company.com
5 again by explicitly typing the URL in a browser window
6 (instead of following a link), he will get a new value
7 for the <identity> field.

8 In many cases, it is highly desirable to know about the
9 identity of the user when he goes from one site in a
10 domain to a second site in another domain. This may be
11 desirable so that a consistent set of information be
12 displayed to the user across the domains, so that a
13 single-sign on scheme be implemented, or simply for the
14 purpose of identifying the common set of users in the
15 two domains.

16 In an embodiment in accordance with the present
17 invention, basic operation of the system follows a
18 scheme in which each of the different domains uses
19 their own user-tracking mechanism. When using cookies,
20 they each set their own independent cookies at the
21 browser. However, they also follow an additional step
22 of coordinating the identity information contained in
23 the cookie with each-other. This coordination allows
24 the user to be tracked across multiple domains.

25 An example embodiment of a system which can be used to
26 implement the cookie sharing mechanism is shown in
27 Figure 2. The user 201 accesses two sites, first site
28 203 and second site 205. The user 201, the first site
29 203, the second site 205 and a cookie coordination

1 database 207 are connected together by the network 209.
2 When the user accesses site 1 203, the site assigns its
3 own identity to the user. When the user accesses the
4 first site 203, the site 203 uses its user tracking
5 mechanism to assign an identity to the user, and stores
6 information about the user at the cookie coordination
7 database 207. The first site 203 also directs the
8 client to access a resource at the second site 205.
9 This can be done by means of a HTTP redirection, or by
10 means of placing a link to the second site 205 in the
11 page being sent to the client by the first site 203.
12 The link or redirection encapsulates information about
13 the location of the record in the cookie coordination
14 database 207 identifying the client information. When
15 the second site 205 is accessed by the client, the site
16 decapsulates the location of the client in the cookie
17 coordination database 207, and creates its own user
18 tracking mechanism to identify the client. The second
19 site 105 can also store information about its user
20 tracking mechanism in the cookie coordination database
21 207 enabling the first site to 203 access the identity
22 of the user at the second site 205.

23 As an example, consider the case where the user
24 tracking mechanism used by the two sites is a cookie.
25 The first site 203 will place a cookie cookie-one in
26 the user's browser. Let us assume that the cookie has
27 an identity field which is selected to have the value
28 of id-one by the first site. The first site 203 stores
29 this information as the k-th record in the database
30 207. It includes a link to an image in the page being
31 sent to the client which asks the client to load an

1 image located at the relative URL /location=k/image.gif
2 at the second site. Since the link directs the client
3 to load an image from the second site, the second site
4 will also place its own independent cookie at the
5 user's browser. Let us say that the cookie contains an
6 identity id-two for the second site 205. The second
7 site 205 can now update the k-th record at the database
8 207 to store the value of id-two. It can also look up
9 the fact that this is the same client as the one
10 identified by id-one at the first site 203.

11 Those skilled in the art will realize that there are
12 other mechanisms to direct the client to the second
13 site. As an example, the well-known HTTP redirection
14 mechanisms using a HTTP response code of 301, 302, 305
15 or 307 can be used to direct the client to second site,
16 and back from the second site to the first site. The
17 URL can encapsulate the location of the record in the
18 database in a different number of ways. Similarly, the
19 information correlating the two cookies can be store
20 individually in the cookies itself instead of the
21 database 207. This allows the database record entry to
22 be removed after the second site has obtained the
23 correlation information. The database 207 can also
24 remove records on a least-recently used bases in order
25 to free up the space, or it can remove a cookie entry
26 after it has been inactive for some time. Since the
27 cookie coordination database 207 servers the purpose of
28 cookie coordination, it can be called a cookie
29 coordinator.

1 The steps involved in the cookie correlation as
2 described in the environment of Figure 2 are outlined
3 in the flowcharts shown in Figure 3 and Figure 4. The
4 steps of Figure 3 are executed by the first web site
5 when a client requests access to a page at the first
6 web site at the initial step of 301. In the next step
7 303, the first web site assigns an identity to the
8 client and stores a client record in the database. In
9 the next step 305, the first web site creates a link
10 for the second site which encapsulates information
11 about the location of the client record in the cookie
12 coordinator database. In the next step 307, the first
13 web site creates an user-tracking mechanism for the
14 user that includes the identity information. This
15 mechanism could be a cookie or a rewritten tag within
16 an uRL. In the step 309, the first web site directs the
17 client to the second web site. The first web site then
18 exits the algorithm in step 311.

19 The second web site executes the steps outlined in
20 Figure 4 when it receives the request from the
21 redirected user. The algorithm is entered in step 401.
22 In the next step 403, the second web site decpasulates
23 the information about the location record for the
24 client in the cookie coordinator. In step 405, the
25 second web site uses the information in the client
26 record accessed from the database in conjunction with
27 its own user tracking mechanism to track the second
28 user. It then exits the algorithm in step 407. The
29 second site can use the same identifier for the user as
30 the first web site, or it can use a different
31 identifier and store the identifier information in the

1 cookie coordinator database. In other cases, the second
2 site can create a third identifier which includes both
3 the identifier used at the first site, and the
4 identifier used at the second site as sub-components,
5 and store the third identifier as part of the user
6 tracking mechanism.

7 In alternate embodiments of the present invention, the
8 coordinated user information can be used in a variety
9 of ways. One of the uses of the coordination
10 information is to share access control and
11 authentication information. As an example, the first
12 web site may have authenticated the credentials of the
13 users and created a cookie with the appropriate
14 credentials. The second site wants to reuse the same
15 credentials instead of asking the user to provide its
16 credentials once again. The credential information can
17 be stored in the cookie coordinator database, and the
18 second site can look up the cookie coordinator database
19 to check for credentials rather than challenging the
20 user once again. This mechanism enables a single
21 sign-on mechanism across the two domains to which two
22 web-servers may belong.

23 Other embodiments employ the cookie coordination
24 mechanism to create personalized pages for an user on
25 the basis of the preferences or characteristics stored
26 by the user at another site. As an example, an user may
27 have stated that he has an interest in sports news when
28 he created a personalized profile for the first web
29 site. When the second web site can correlate its
30 cookies with the cookies of the first web site, it can

1 infer that the user is interested in sports news, and
2 create pages incorporating sports news even though the
3 user did not provide this information to the second web
4 site. Thus, sharing of cookie information can lead to
5 sharing of user preferences and other information
6 across multiple domains.

7 In additional alternate embodiments of the present
8 invention, each of the servers in different domains
9 can maintain a private cookie at the browser; with each
10 web server accessing the cookie coordinator when the
11 private cookie it maintains is received by a
12 web-server; and the cookie coordinator maps the
13 identities contained in the cookies from different net
14 domains to a single identity common across the multiple
15 domains. In some cases, the single identity is stored
16 in the private cookie maintained by the server in the
17 domain.

18 In some of these additional alternate embodiments of
19 the present invention, the embodiment may use a single
20 identity for the users across the different domains.
21 While each private cookie established in each domain
22 contains a different identity, the cookie coordinator
23 maintains a single identity which is used to correlate
24 information from the different clients. The cookie
25 coordinator learns the mapping of the various
26 identities placed in each private cookie, and learns
27 the mapping of the identities placed in the private
28 cookie to the single identity.

1 An additional alternate embodiments of the present
2 invention, includes an apparatus shown in Figure 5.
3 The apparatus in Figure 5 includes: a web server
4 interface to interface with a first web server in a
5 first DNS domain 510, and a second web server in a
6 second DNS domain 520, wherein the first web server
7 uses a first user tracker 512 to collect client
8 information and stores the client information as a
9 client record in a cookie coordinator database 560; a
10 redirector 530 for the first web server directing a
11 client to access a resource at the second web server;
12 an encapsulator 514 for said resource encapsulating
13 information about a location of the client record in
14 the database; a decapsulator 540 for the second web
15 server decapsulating the location and retrieving the
16 client record from the database 560; and a second user
17 tracker 550 for the second web server using the client
18 record in conjunction with a second user tracking
19 mechanism.

20 The present invention can be realized in hardware,
21 software, or a combination of hardware and software. A
22 visualization tool according to the present invention
23 can be realized in a centralized fashion in one
24 computer system, or in a distributed fashion where
25 different elements are spread across several
26 interconnected computer systems. Any kind of computer
27 system - or other apparatus adapted for carrying out
28 the methods and/or functions described herein - is
29 suitable. A typical combination of hardware and
30 software could be a general purpose computer system
31 with a computer program that, when being loaded and

1 executed, controls the computer system such that it
2 carries out the methods described herein. The present
3 invention can also be embedded in a computer program
4 product, which comprises all the features enabling the
5 implementation of the methods described herein, and
6 which - when loaded in a computer system - is able to
7 carry out these methods.

8 Computer program means or computer program in the
9 present context include any expression, in any
10 language, code or notation, of a set of instructions
11 intended to cause a system having an information
12 processing capability to perform a particular function
13 either directly or after either or both of the
14 following conversion to another language, code or
15 notation, and/or reproduction in a different material
16 form.

17 Thus the invention includes an article of manufacture
18 which comprises a computer usable medium having
19 computer readable program code means embodied therein
20 for causing a function described above. The computer
21 readable program code means in the article of
22 manufacture comprises computer readable program code
23 means for causing a computer to effect the steps of a
24 method of this invention. Similarly, the present
25 invention may be implemented as a computer program
26 product comprising a computer usable medium having
27 computer readable program code means embodied therein
28 for causing a function described above. The computer
29 readable program code means in the computer program
30 product comprising computer readable program code means

1 for causing a computer to effect one or more functions
2 of this invention. Furthermore, the present invention
3 may be implemented as a program storage device readable
4 by machine, tangibly embodying a program of
5 instructions executable by the machine to perform
6 method steps for causing one or more functions of this
7 invention.

8 It is noted that the foregoing has outlined some of the
9 more pertinent objects and embodiments of the present
10 invention. This invention may be used for many
11 applications. Thus, although the description is made
12 for particular arrangements and methods, the intent and
13 concept of the invention is suitable and applicable to
14 other arrangements and applications. It will be clear
15 to those skilled in the art that modifications to the
16 disclosed embodiments can be effected without departing
17 from the spirit and scope of the invention. The
18 described embodiments ought to be construed to be
19 merely illustrative of some of the more prominent
20 features and applications of the invention. Other
21 beneficial results can be realized by applying the
22 disclosed invention in a different manner or modifying
23 the invention in ways known to those familiar with the
24 art.